

Weekly Report

September 3, 2017

1 Work

- 本周周按照Review意见将文章中的细节部分进行修改，主要增加了数学符号表格，定义了某些术语和变量的含义，基本完成cover letter。因为时间关系，还有几个点没有完成，下周会重新整理一下分析任务，尽快完成。
- 本周对一些向量化表达的文章进行了简单的分类归纳。暂时只收录了基于或者类似于word2vec的文章，还有基于其他方法的文章如CNN和RNN没有加入，后期可以根据方法进行分类，比如监督学习无监督学习。

1.1 2Vec综述

1.2 工作进度

Table 1: 综述

数据类型	文章名
Text	<ol style="list-style-type: none"> 1. Distributed Representations of Words and Phrases and their Compositionality (word2vec) 2. Distributed representations of sentences and documents (doc2vec) 3. Novel2Vec: Characterising 19th Century Fiction via Word Embeddings 4. cite2vec: Citation-Driven Document Exploration via Word Embeddings 5. Discovery of Evolving Semantics through Dynamic Word Embedding Learning 6. Mixing Dirichlet Topic Models and Word Embeddings to Make lda2vec 7. sense2vec - a fast and accurate method for word sense disambiguation in neural word embeddings 8. emoji2vec: Learning Emoji Representations from their Description
Network	<ol style="list-style-type: none"> 1. DeepWalk: Online Learning of Social Representations 2. node2vec: Scalable Feature Learning for Networks 3. struc2vec: Learning Node Representations from Structural Identity 4. struc2vec: Learning Node Representations from Structural Identity
Trajectory	<ol style="list-style-type: none"> 1. Exploring the Context of Locations for Personalized Location Recommendations 2. POI2Vec: Geographical Latent Representation for Predicting Future Visitors 3. A General Multi-Context Embedding Model for Mining Human Trajectory Data 4. From Word Embeddings to Item Recommendation 5. A Data Grouping CNN Algorithm for Short-Term Traffic Flow Forecasting
Video	<ol style="list-style-type: none"> 1. Video2Vec: Learning Semantic Spatial-Temporal Embeddings for Video Representation 2. Learning Temporal Embeddings for Complex Video Analysis
Item	<ol style="list-style-type: none"> 1. item2vec: neural item embedding for collaborative filtering

Table 3: 工作进度

任务	当前进度	截止时间
location2vec修改	本周周按照Review意见将文章中的细节部分进行修改，还有几个点没有完成，下周会重新整理一下分析任务，尽快完成。	8月30日
投影		9月30日
*2Vec综述	本周对目前收集的文章进行分类归纳，下周可以扩大一下文章的方法类型。	10月30日
专利	已经联系过律师，会尽快完成。	

2 Paper Reading

2.1 Illustration2Vec: A Semantic Vector Representation of Illustrations

利用CNN，将网络上的图片和它对应的tag进行训练，得到一种向量表达。

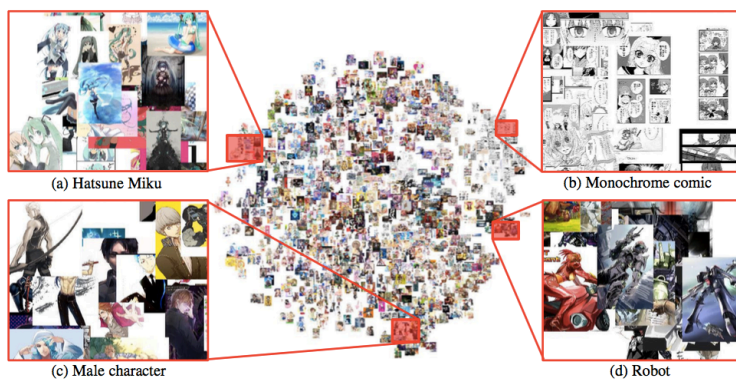


Figure 1: Illustration2Vec

2.2 MobiSeg: Interactive Region Segmentation Using Heterogeneous Mobility Data

本文设计了一种功能区分块的可视分析系统，基于非负矩阵分解的方法对地理区域进行聚类，用户可以通过交互干预分块的结果。

2.3 emoji2vec: Learning Emoji Representations from their Description

文章基于对emoji的描述文字（单词已经被嵌入到向量空间）学习得到相对应的emoji的特征向量。

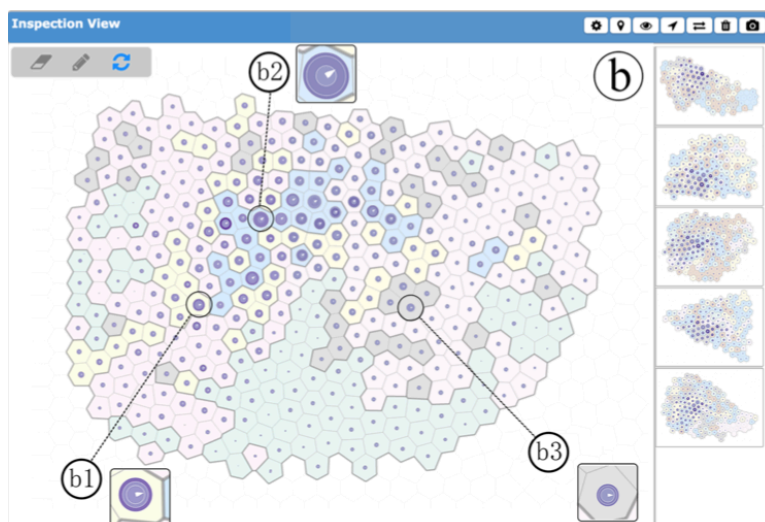


Figure 2: MobiSeg

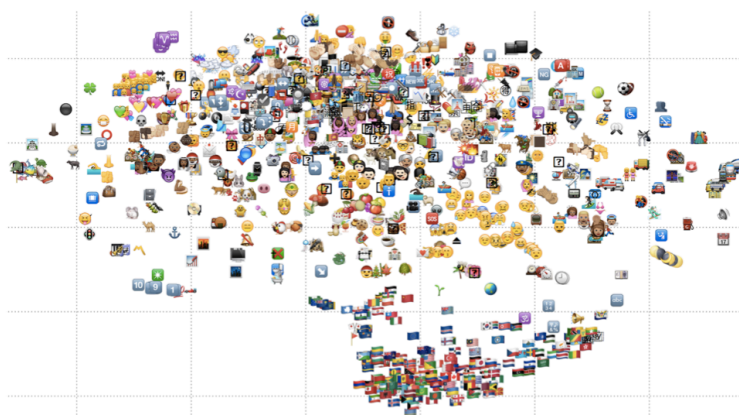


Figure 3: emoji2vec